Signal Prediction With Input Identification

Jer-Nan Juang
NASA Langley Research Center, Hampton, Virginia

Ya-Chin Chen
Institute for Computer Applications in Science and Engineering (ICASE), Hampton, Virginia

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the lead center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart and peer-reviewed formal professional papers, but having less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

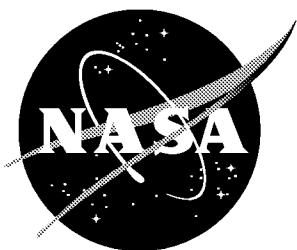- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results... even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at *http://www.sti.nasa.gov*

- E-mail your question via the Internet to help@sti.nasa.gov

- Fax your question to the NASA STI Help Desk at (301) 621–0134

- Phone the NASA STI Help Desk at (301) 621–0390

- Write to:
  NASA STI Help Desk
  NASA Center for AeroSpace Information
  7121 Standard Drive
  Hanover, MD 21076–1320

NASA/TM–1999–209705

Signal Prediction With Input Identification

Jer-Nan Juang
NASA Langley Research Center, Hampton, Virginia

Ya-Chin Chen
Institute for Computer Applications in Science and Engineering (ICASE), Hampton, Virginia

October 1999

Signal Prediction With Input Identification

Jer-Nan Juang *
NASA Langley Research Center
Hampton, VA 23681

Ya-Chin Chen[†]
ICASE/NASA Langley Research Center
Hampton, VA 23681.

Abstract

A novel coding technique is presented for signal prediction with applications including speech coding, system identification, and estimation of input excitation. The approach is based on the blind equalization method for speech signal processing in conjunction with the geometric subspace projection theory to formulate the basic prediction equation. The speech-coding problem is often divided into two parts, a linear prediction model and excitation input. The parameter coefficients of the linear predictor and the input excitation are solved simultaneously and recursively by a conventional recursive least-squares algorithm. The excitation input is computed by coding all possible outcomes into a binary codebook. The coefficients of the linear predictor and excitation, and the index of the codebook can then be used to represent the signal. In addition, a variable-frame concept is proposed to block the same excitation signal in sequence in order to reduce the storage size and increase the transmission rate. The results of this work can be easily extended to the problem of disturbance identification. The basic principles are outlined in this report and differences from other existing methods are discussed. Simulations are included to demonstrate the proposed method.

1 INTRODUCTION

In the past decade, a number of advanced technologies have been employed to represent speech signals digitally for use in communication-related operations such as audio transmission, storage, manipulation, speech recognition, and even speech synthesis. These operations can be performed more efficiently by reducing the amount of information needed to represent a given speech signal. The term "speech coding", or simply "coding" is thus introduced in speech processing. In speech coding, a major objective is to represent the digital signal

with as few bits as possible, i.e., to compress the signal. The degree of compression depends on the cost of transmission or storage, the cost of coding the digital speech signal, and speech quality requirements. Before 1980, the high cost of coding and low speech quality made the speech coding impractical. However, with the improved digital signal processing hardware capability and significant progress in speech coding research, speech coding is now widely used in a variety of applications. Speech coding techniques proposed and developed over the past decade can be divided into two general categories; waveform coders and voice coders (vocoders) [1].

In most of the waveform coding techniques, the samples are processed by the scalar quantization. A scalar quantizer operates on a single sample at a time and represents each sample by a sequence of levels through a mapping function. The output of the quantizer, namely the quantized signal, can hence be coded by binary digits. On the other hand, a block of samples may be quantized as a single entity through a mapping function, which is called vector quantization.

In contrast with the waveform coding, voice coding divides the speech problem into two parts; part one creates an analytical model of the vocal tract, and part two synthesize an analytical representation of the input excitation. The true input is never measured but the idea is to reconstruct the recorded signal by convolving the analytical model with a synthesized input. Typically, the analytical model structure is assumed to have all poles and the synthesized input is assumed to be a periodic impulse train with period equal to the fundamental frequency. For unvoiced speech, the excitation is a white noise sequence [1, 2].

Linear Predictive Coding (LPC or LP) is a voice coding approach widely used in practice today. The objectives of LP analysis are to estimate the coefficients of an all-pole model representing the vocal tract, to determine analytically the type of excitation, and to estimate the fundamental frequency, and its gain coefficients. Different LPC-type speech analysis and synthesis schemes differ primarily in the type of input signal which is generated for speech synthesis. Several schemes have been proposed for generating the input signal; residual excited linear prediction (RELP) vocoder, multipulse LPC vocoder, code-excited linear prediction (CELP) [3] and vector sum excited linear prediction (VSELP) [4].

The advanced speech coders since the 1990s are based on the LPC scheme using vector quantization (VQ). In the LPC-type vocoder, the bulk of the transmission rate is used to transfer the synthesized excitation sequence. Therefore, how to synthesize excitation

2

efficiently and effectively becomes very important. In [5], the vector quantization introduced includes codebook coding, tree coding, and trellis coding. However, as most of the coders use codebook coding, this method is of particular interest. In codebook coding, the set of possible output sequences is arranged in a codebook whose elements are not restricted in any way. When the optimum output sequence is searched, the corresponding index of that sequence is transmitted. In fact, codebook coding is impractical when the size of the codebook is large. Some effort for searching the optimum sequence has been done [6], for example the binary tree search. The proper design of the codebook is the key to a success for LPC-type speech coding.

The conventional blind equalization is aimed at recovering the input signal applied to a linear time-invariant system from the observed signal developed at the output of the system [7]. In other words, blind equalization is a special kind of adaptive inverse filtering that operates without access to the source of the input signal. In digital communications, the input signal is commonly called the transmitted signal. The time-invariant system is referred to as the channel. There are two general approaches developed to achieve this task; the Constant Modulus Algorithm (CMA) [7, 8] and Decision Directed (DD) [9, 10] equalizer. The main idea is to keep the output of the equalizer at constant modulus (absolute value) [7, 8, 11]. The input signal will have some known property, which helps determine how the observed signal has been corrupted. In [12, 13], the blind adaptive prediction exploited the constant modulus property to keep the prediction error at each estimate within predefined bounds. The use of constant modulus is to modulate the prediction error to a constant value. Once the prediction error is modulated to a sufficiently small value, the prediction derived from blind equalization becomes reliable. Applying the idea of constant modulus to speech coding enables the excitation quantized at the same time while the LPC coefficients are updated. *Note that the equalizer output is the recovered input signal that is similar to the LPC-based quantized input.* However, this leads an open question of how small the given modulus needs to be selected [14]. Furthermore, it will be shown in this paper that the blind adaptive prediction, which only contains a unit modulus, is unable to obtain the steady LPC solution for speech coding. A natural alternative is to expand then single modulus to become multiple modulus. As a result, the *adaptive multi-modulus blind predictor* is proposed in this paper.

In [15], the multiple modulus concept was proposed to deal with blind equalization of

3

signal, such as Multilevel Quadrature-Amplitude Modulation (M-ary QAM). The major difference is that, contrary to the approaches presented in [15, 16], our proposed *adaptive multi-modulus blind predictor* does not specify a priori the modulus. Moreover, our approach combines recursive least-squares with DD approaches to determine the modulus recursively [17]. In contrast, the first proposed MMA (multiple modulus algorithm) [15], uses a straightforward generalization of the CMA cost function to derive its update and the second one, DAMA (decision adjusted modulus algorithm), is a hybrid of the CMA and the DD approaches.

In this paper, we propose a novel coding technique for speech compression. The technique does not require separate solutions for the equalizer coefficients and input quantization. The equalizer coefficients in this paper are nothing but the coefficients of an all-pole model. The approach is to integrate the input identification into the adaptive estimation of the equalizer coefficients. The goal is to make the proposed technique feasible for real-time implementation in practice. The estimation of equalizer coefficients and the input identification are obtained recursively by the coefficient smoothing technique [17]. The input signal is generated without using a separate quantization scheme when the predictor is updated. The input codebook is derived analytically instead of generating it based on the stochastic assumptions [18]. After the entire process completed, the parameters to be quantized before transmission or stored are the coefficients, the gains of the input, and the index of the input sequence. The geometry space concepts lead to an intensive and complete explanation of the proposed technique.

Regarding the aim of low bit rate coding, the conventional coders usually deal with the speech by frames of samples. However, it is likely that this may not be the best way to describe the non-stationary behavior of the sound sources. On the other hand, the precise coding based on sample-by-sample can always produce high quality with negligible coding distortion and negligible coding delay. The blind adaptive prediction is originally proposed to perform the prediction on the sample-by-sample to overcome the problems that the conventional LP model suffers from the stationary assumptions involved. As a result, the approach that proposed in this paper includes a variable frame concept for transmission and storage. The sample-by-sample scheme is used to do precise coding and obtain the high quality. The same excitation signal is then blocked into frames to be transmitted or stored. Hence, the resulting bit rate can be reduced.

2 ALGORITHM

This section begins with a brief description of conventional linear predictive coding and blind equalization. The linear predictive techniques were developed mainly for speech coding, whereas the blind equalization techniques were derived for input identification, i.e., transmitted signal recovery for digital communications.

2.1 Linear Prediction (LP)

Linear prediction techniques were first used for speech analysis and synthesis by Itakura and Saito [19], and Atal and Schroeder [20], which foster further work in coding, recognition, enhancement and so on [2, 21, 22, 23]. A general flowchart of the LP modeling is shown in figure 1, and the predicted value is defined by
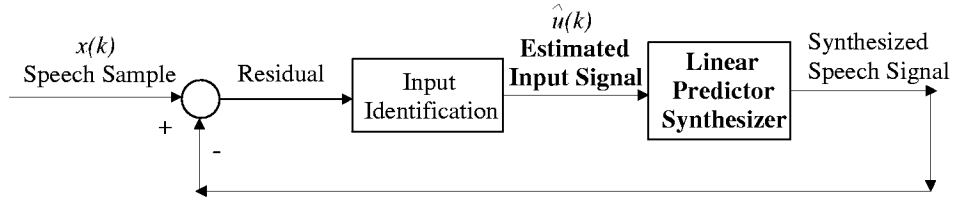

Figure 1: LPC modeling diagram

$$\widehat{x}(k) + \sum_{i=1}^{n} \theta_i \widehat{x}(k-i) = \widehat{u}(k) \tag{1}$$

where $\widehat{x}(k)$ is the synthesized speech signal and $\widehat{u}(k)$ is its quantized input signal at time k. Equation (1) shows that the synthesized signal $\widehat{x}(k)$ is a linear function of the current input signal $\widehat{u}(k)$ and its past signal $\widehat{x}(k-i)$ weighted by the tap constant value θ_i for $i = 1, 2, \ldots, n$ where n is an integer greater than zero. In signal processing, Eq. (1) is called the *closed-loop formulation* for computing the synthesized signal $\widehat{x}(k)$. The tap weights θ_i are commonly called the LP coefficients that constitutes an all-pole model. The LP coefficients θ_i for $i = 1, 2, \ldots, n$ and the quantized input signal $\widehat{u}(k)$ for $k = 1, 2, \ldots, \ell$ with ℓ being the length of data length may be obtained by minimizing the squared error between real and synthesized speech signals, which is defined by

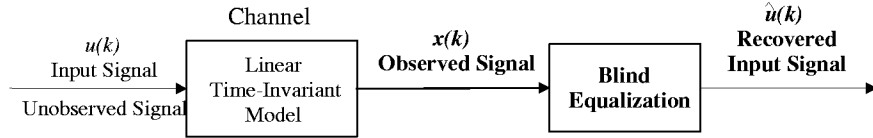$$J = E\{| x(k) - \widehat{x}(k) |^2\} \tag{2}$$

Figure 2: Blind Equalization

where $E\{\cdot\}$ denotes the expectation operator. Accordingly, a number of methods such as the autocorrelation method, the covariance method, the lattice method, and so forth [24] were so developed as the formulations for easrching LP solutions.

2.2   Blind Equalization

Blind Equalization in digital communication is a deconvolution process to recover the input signal applied to a time-invariant system. It is a special kind of adaptive inverse filtering that operate without access the transmitted signal (i.e., input signal). The name "blind equalization" refers to the ability of an adaptive algorithm to perform deconvolution in a *blindfolded* or *self-recovered fashion*. Figure 2 shows a general flowchart where the channel includes the combined effects of a transmit filter, a transmission medium, and a receiver filter that may be represented by a linear time-invariant system or a linear predictor. The objective is to estimate the coefficients of the linear predictor and the input signal, given the observed signal $\hat{x}(k)$ for $k = 1, 2 \ldots, \ell$. The input signal is generally in the form of binary sequence.

In [12, 13], a procedure was proposed for estimation of the LP coefficients. It was aimed at overcoming the difficulties inherent in the non-stationarity of the signals to be modeled. The approach is to keep the prediction error of representation at each estimate to within a predefined set of bounds rather than minimize the mean squared error in Eq. (2). The output of the equalizer (i.e., the recovered input signal) can be obtained by

$$\widehat{u}(k) = \sum_{i=0}^{n} \theta_i(k) x(k-i) \tag{3}$$

where $\theta_i(k)$ at each time step k for $i = 0, \ldots, n$ are coefficients to be determined. Note that Eq. (3) uses the observed quantity x rather than the synthesized signal $\hat{x}$ in Eq. (1). Equation (3) is commonly called the open-loop formulation in signal processing whereas Eq. (1) is referred to as the closed-loop formulation.

6

In 1980, Godard proposed a family of constant modulus blind equalization algorithms for use in two-dimensional digital communication systems [7]. Among them, the so called CMA (constant modulus algorithm) is derived from the cost function as

$$J_{CMA} = E\{(\widehat{u}(k)^2 - 1)^2\} \tag{4}$$

Once the equalization is achieved, the output of the equalizer is modulated to approach to ± 1. The LP model from Eq. (3) can be modified by normalizing the coefficients $\theta_i(k)$ at each time step k for $i = 0, \ldots, n$ such that $\theta_0(k) = 1$ to yield

$$\widehat{u}(k) = x(k) + \sum_{i=1}^{n} \theta_i(k)x(k-i) \tag{5}$$

The coefficient normalization is equivalent to scaling the observed signal $x(k)$ for $k = 1, 2, \ldots, \ell$ such that its absolute maximum value is unity. Equation (5) is identical to the LP model shown in Eq. (1) except that it is an *open-loop formulation* which will be explained later.

Using the blind equalization method, the LP modeling problem is formulated as a constrained optimization problem

$$\min \quad \|\boldsymbol{\Theta}(k+1) - \boldsymbol{\Theta}(k)\|_2^2$$
$$\text{subject to} \quad |x(k) + \mathbf{X}^T(k-1)\boldsymbol{\Theta}(k)| = \epsilon$$

where ϵ is a predefined positive constant value. The coefficient vector $\boldsymbol{\Theta}(k)$ contains the coefficients of the equalizer,

$$\boldsymbol{\Theta}(k) = [\theta_1(k) \ \theta_2(k) \ldots \theta_n(k)]^T$$

and $\mathbf{X}(k-1)$ is an observed sequence,

$$\mathbf{X}(k-1) = [x(k-1) \ x(k-2) \ \ldots \ x(k-n)]^T$$

This constrained optimization problem was reformulated in [12, 13] to a Lagrange equivalent and the stochastic gradient descent strategy was adopted to solve for the LP coefficient vector. These algorithms need iterations between each consecutive sample to ensure that the constraints are satisfied. The prediction implemented to illustrate the performance was open-loop prediction. Although it is not a closed-loop prediction as required in speech

7

coding, the prediction algorithm may provide a new solution for input identification and quantization because of the predefined constant value of the equalizer output. Indeed, the definition of the equalizer output in Eq. (5) may be considered as the binary input of the LPC-type vocoders.

2.3 Adaptive Multi-Modulus Blind Equalization

To make the blind equalization algorithm applicable for speech coding (i.e., closed-loop prediction), the conventional algorithm must be modified and advanced. A new approach is derived in this section. The undefined value ϵ will be determined autonomously and adaptively rather than randomly picked. The input signal is not restricted to a predefined single constant value commonly used in the blind adaptive prediction (see Eq. (3)).

Let the values of the output of the equalizer from Eq. (5) be the series sum of a binary stream multiplied by scalars,

$$\widehat{u}(k) = -\sum_{i=1}^{N} \phi_i(k)\delta_i(k) \tag{6}$$

where $\delta_i(k) \in \{\pm 1\}$ is a binary stream and $\phi_i(k)$ is a weighting coefficient. The goal is to make the weighting coefficient $\phi_i(k)$ invariant with respect to k. Substituting Eq. (6) into Eq. (5) yields

$$
\begin{aligned}
x(k) + \mathbf{X}^T(k-1)\boldsymbol{\Theta}(k) &= -\sum_{i=1}^{N} \delta_i(k)\phi_i(k) \\
&= -\boldsymbol{\Delta}^T(k)\boldsymbol{\Phi}(k) \tag{7}
\end{aligned}
$$

where $\boldsymbol{\Delta}(k) = [\delta_1(k)\ \delta_2(k)\ ...\ \delta_N(k)]^T$ and $\boldsymbol{\Phi}(k) = [\phi_1(k)\ \phi_2(k)\ ...\ \phi_N(k)]^T$. Equation (7) can be rearranged to become

$$
\begin{aligned}
x(k) &= -\mathbf{X}^T(k-1)\boldsymbol{\Theta}(k) - \boldsymbol{\Delta}^T(k)\boldsymbol{\Phi}(k) \\
&= -\left[\ \mathbf{X}^T(k-1)\ \ \boldsymbol{\Delta}^T(k)\ \right]\left[\begin{array}{c} \boldsymbol{\Theta}(k) \\ \boldsymbol{\Phi}(k) \end{array}\right] \tag{8}
\end{aligned}
$$

Define two new quantities

$$
\begin{aligned}
\mathbf{V}(k-1) &= \left[\begin{array}{c} \mathbf{X}(k-1) \\ \boldsymbol{\Delta}(k) \end{array}\right] \\
\boldsymbol{\Psi}(k) &= \left[\begin{array}{c} \boldsymbol{\Theta}(k) \\ \boldsymbol{\Phi}(k) \end{array}\right] \tag{9}
\end{aligned}
$$

Equation (8) becomes

$$x(k) = -\mathbf{V}^T(k-1)\mathbf{\Psi}(k) \tag{10}$$

In practice for speech coding, the coefficient vector $\mathbf{\Psi}(k)$ should be constrained such that

$$\mathbf{\Psi}(k+1) = \mathbf{\Psi}(k) = \mathbf{\Psi} \tag{11}$$

for all k, i.e., independent of k. Imposing the constraint, Eq. (10) becomes

$$x(k) = -\mathbf{V}^T(k-1)\mathbf{\Psi} \tag{12}$$

The gain coefficient vector, $\mathbf{\Phi}(k)$, is now integrated into the newly defined coefficient vector $\mathbf{\Psi}$ and can be seen as a deterministic gain vector of the input excitation $\widehat{u}$.

Schroeder and Atal [3] used a Gaussian codebook to encode the input. After examining the first-order cumulative probability distribution function for the prediction residual, they found it resembled a corresponding Gaussian distribution function with the same mean and variance. In contrast, our prediction residual from the blind equalization is only a binary stream, $\mathbf{\Delta}(k)$. A binary codebook may be introduced to determine and encode the input. Binary coding is less restrictive than the conventional Gaussian codebook for generating the input, because it does not impose any assumption on the stochastic process.

It is quite simple to generate the binary codebook. For example, a 2-bit codebook consists of 4 code vectors as follows

$$\begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{pmatrix}$$

and a 3-bit codebook has the form

$$\begin{pmatrix} 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \end{pmatrix}$$

An N-bit codebook requires a collection of 2^N codes which can be used in Eq. (10) to estimate $\widehat{u}(k)$. By selecting the minimum error, which results from the difference between the 2^N estimated values and the original signal, the optimum N-bit input binary sequence can be determined. Its corresponding index, which is in a binary format using N- bit, can then be stored or transmitted.

To solve Eq. (12) for the constant vector $\boldsymbol{\Psi}$ is quite simple assuming that the optimum N-bit input binary sequence $\boldsymbol{\Delta}(k)$ for all k is known a priori. Defining the quantities

$$\mathbf{X} = \begin{bmatrix} x(n+1) & x(n+2) & \cdots & x(\ell) \end{bmatrix}^T \tag{13}$$

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}(n) & \mathbf{V}(n+1) & \cdots & \mathbf{V}(\ell-1) \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{X}(n) & \mathbf{X}(n+1) & \cdots & \mathbf{X}(\ell-1) \\ \boldsymbol{\Delta}(n+1) & \boldsymbol{\Delta}(n+2) & \cdots & \boldsymbol{\Delta}(\ell) \end{bmatrix} \tag{14}$$

Equation (12) thus produces

$$\mathbf{X} = -\mathbf{V}^T \boldsymbol{\Psi} \tag{15}$$

The quantity $\mathbf{X}$ is a $(\ell-n)\times 1$ vector, $\boldsymbol{\Psi}$ is an $(n+N)\times 1$ vector, and $\mathbf{V}^T$ is a $(\ell-n)\times(n+N)$ matrix. *From Eq. (15), there exits a solution for $\boldsymbol{\Psi}$ if and only if the vector $\mathbf{X}$ is in the column space generated by the columns of $\mathbf{V}^T$.* For the case where $(\ell - n) > (n + N)$ (i.e., more equations than unknowns), it is generally impossible to satisfy such a sufficient and necessary condition unless the signal to be synthesized is generated from a noise-free finite-dimensional linear system. The optimum solution is then the least-squares solution, i.e.,

$$\widehat{\boldsymbol{\Psi}} = -(\mathbf{V}^T)^\dagger \mathbf{X} \tag{16}$$

where $\dagger$ means the pseudo-inverse and $\widehat{\boldsymbol{\Psi}}$ implies the estimated quantity of $\boldsymbol{\Psi}$, that is,

$$\widehat{\boldsymbol{\Psi}} = \begin{bmatrix} \widehat{\boldsymbol{\Theta}}(k) \\ \widehat{\boldsymbol{\Phi}}(k) \end{bmatrix} \tag{17}$$

for any k. The least-squares solution minimizes the equation error between the real signal $\mathbf{X}$ and the estimated signal $\widehat{\mathbf{X}}$, i.e.,

$$\widehat{\mathbf{X}} = -\mathbf{V}^T \widehat{\boldsymbol{\Psi}} \tag{18}$$

Note that the quantity $\widehat{\mathbf{X}}$ is an open-loop estimation and thus not a synthsized speech signal.

One problem remains to be solved. The optimum N-bit input binary sequence

$$\boldsymbol{\Delta} = \begin{bmatrix} \boldsymbol{\Delta}(n+1) & \boldsymbol{\Delta}(n+2) & \cdots & \boldsymbol{\Delta}(\ell) \end{bmatrix}$$

needs to be determined. Given any sequence, say

$$\boldsymbol{\Delta}_i = \begin{bmatrix} \boldsymbol{\Delta}_i(n+1) & \boldsymbol{\Delta}_i(n+2) & \cdots & \boldsymbol{\Delta}_i(\ell) \end{bmatrix}$$

10

let us denote the least-squares prediction error to be

$$Err_i = \mathbf{X} - \widehat{\mathbf{X}}_i = \mathbf{X} + \mathbf{V}_i^T \widehat{\mathbf{\Psi}}_i \tag{19}$$

where

$$\widehat{\mathbf{X}}_i = -\mathbf{V}_i^T \widehat{\mathbf{\Psi}}_i \tag{20}$$

and

$$
\begin{aligned}
\widehat{\mathbf{\Psi}}_i &= -(\mathbf{V}_i^T)^\dagger \mathbf{X} \\
\mathbf{V}_i &= \left[\begin{array}{cccc} \mathbf{X}(n) & \mathbf{X}(n+1) & \cdots & \mathbf{X}(\ell-1) \\ \mathbf{\Delta}_i(n+1) & \mathbf{\Delta}_i(n+2) & \cdots & \mathbf{\Delta}_i(\ell) \end{array} \right]
\end{aligned} \tag{21}
$$

Among all possible choices of $\mathbf{\Delta}_i$ for $i = 1, \ldots, (2^N)^l$, one should pick the one that minimizes the norm of Err_i. There are several ways of determining the optimum coefficient vector $\widehat{\mathbf{\Psi}}$ and the optimum N-bit input binary sequence $\mathbf{\Delta}$ from Eq. (16). The key idea is to choose the N-bit input binary sequence such that the columns of $\mathbf{V}^T$ generate a column space to include the vector $\mathbf{X}$ as much as possible.

Here we introduces a recursive least-squares technique that minimizes the cost function defined as

$$\min \| \sum_{i=1}^{k} \lambda^{k-i} [x(i) + \mathbf{V}^T(i-1)\mathbf{\Psi}(k)]^2 \| \tag{22}$$

where $\mathbf{\Psi}$ is the smoothing coefficient vector and $0 < \lambda \leq 1$ is the forgetting factor weighting the data. The most recent data is given unit weight, but data that is n time steps old is weighted by λ^n. The method is commonly called exponential forgetting.

The recursive least-squares algorithm is summarized in the following. At the time index k, choose the N-bit input $\mathbf{\Delta}(k)$ among all possible binary combinations such that the estimation error is minimum, i.e.,

$$
\begin{aligned}
\widehat{x}(k) &= -\mathbf{V}^T(k-1)\widehat{\mathbf{\Psi}}(k-1) \\
&= -\left[\begin{array}{cc} \mathbf{X}^T(k-1) & \mathbf{\Delta}^T(k) \end{array} \right] \left[\begin{array}{c} \widehat{\mathbf{\Theta}}(k-1) \\ \widehat{\mathbf{\Phi}}(k-1) \end{array} \right]
\end{aligned} \tag{23}
$$

$$e_{\min}(k) = \min_{\mathbf{\Delta}(k)} \| x(k) - \widehat{x}(k) \| \tag{24}$$

Both the coefficient vector $\widehat{\mathbf{\Theta}}$ of the blind predictor and the gain coefficient vector $\widehat{\mathbf{\Phi}}$ of the input can be updated recursively by

$$\mathbf{G}(k) = \frac{\mathbf{P}(k-1)\mathbf{V}(k-1)}{\lambda + \mathbf{V}^T(k-1)\mathbf{P}(k-1)\mathbf{V}(k-1)} \tag{25}$$

11

$$\mathbf{P}(k) = \frac{\mathbf{P}(k-1)\left[\mathbf{I} - \mathbf{G}(k)\mathbf{V}^T(k)\right]}{\lambda} \qquad (26)$$

$$\widehat{\mathbf{\Psi}}(k) = \widehat{\mathbf{\Psi}}(k-1) + e_{\min}(k)\mathbf{G}(k) \qquad (27)$$

where $\mathbf{G}(k)$ is the update gain determined by the matrix $\mathbf{P}(k-1)$, the vector $\mathbf{X}(k-1)$, and the scalar λ. The initial values of $\mathbf{P}(0)$ and $\widehat{\mathbf{\Psi}}(0)$ can be arbitrarily assigned. Conventionally, $\mathbf{P}(0)$ and $\widehat{\mathbf{\Theta}}_s(0)$ are assigned as $d\mathbf{I}_{n+N}$ and $\mathbf{0}_{(n+N)\times 1}$, respectively, where d is a large positive number, $\mathbf{I}_{n+N}$ is an identity matrix of dimension $(n+N) \times (n+N)$, and $\mathbf{0}_{(n+N)\times 1}$ is a zero matrix of dimension $(n+N) \times 1$. The estimated coefficient vector, $\widehat{\mathbf{\Psi}}$, is the converged coefficient vector at $k = l$, that is,

$$\widehat{\mathbf{\Psi}} = \widehat{\mathbf{\Psi}}(l) \qquad (28)$$

It is known from the recursive least-squares algorithm that the initialization introduces a *bias* into the parameter estimate $\widehat{\mathbf{\Psi}}$ produced by the recursive least-squares method. For large data lengths, the exact value of the initialization constant is not important. It is noted that some accuracy may be lost when a least-squares problem is solved using the classical approach as described in this section. The reason is that the input and output data are squared to compute the data correlation. There is another method based on orthogonal transformation to avoid the computation of data correlation for the least-squares estimates. The method is commonly called a square root method [27, 28], because it works with the square root of the data correlation.

In the conventional linear predictive coding, the coefficients $\mathbf{\Theta}$ are computed alone using the open-loop formulation via the auto-correlation method or covariance method [1, 25] to minimize the prediction error. Other technology may use the cepstrum analysis to obtain the prediction coefficients, for example the homomorphic, but the computational complexity is a considerable problem in practice. After the coefficients $\mathbf{\Theta}$ are obtained, the input can then be modeled either using vector quantization or any other methods to quantize the prediction error, but most methods involve a priori assumption about the type of the stochastic process. In contrast, our proposed method updates the equalizer coefficients $\mathbf{\Theta}$ together with the gain coefficients $\mathbf{\Phi}$ of the input simultaneously. The application of constant modulus property to constrain the prediction error can be seen as a novel quantization methodology of the input. The input codebook can thus be derived directly from the analytical analysis without involving any assumption about the stochastic properties of the prediction error.

2.3.1 Open-loop prediction

Given the estimated coefficient vector $\widehat{\boldsymbol{\Psi}}$ determined from equation (27), the open-loop prediction equation similar to equation (23) is

$$
\begin{aligned}
\widehat{x}(k) &= -\mathbf{V}^T(k-1)\widehat{\boldsymbol{\Psi}} \\
&= -\left[\ \mathbf{X}^T(k-1)\quad \boldsymbol{\Delta}^T(k)\ \right]\left[\begin{array}{c} \widehat{\boldsymbol{\Theta}} \\ \widehat{\boldsymbol{\Phi}} \end{array}\right]
\end{aligned} \tag{29}
$$

or equivalently

$$
\widehat{x}(k) = -\sum_{i=1}^{n}\widehat{\theta}_i x(k-i) - \sum_{i=1}^{N}\widehat{\phi}_i \delta_i(k) \tag{30}
$$

where $\widehat{\theta}_i$ and $\widehat{\phi}_i$ are constant quantities. The open-loop prediction uses the observed signal to compute the signal prediction. The predicted value $\widehat{x}(k)$ is computed using the past observed signal $x(k-1),\ldots,x(k-n)$ and the input $\widehat{u}(k)$. The binary sequence $\boldsymbol{\Delta}(k) = [\delta_1(k)\ \delta_2(k)\ \ldots\ \delta_N(k)]^T$ of dimension $N \times 1$ is chosen from the N-bit codebook of dimension $N \times 2^N$ with a total of 2^N different codevectors in the codebook. At each time k, only an index from the 2^N possible choices is stored. From hereon, Eq. (30) is used as a predictive model of the adaptive multi-modulus blind equalization for open-loop prediction.

For the case where $N = 1$, Eq. (30) becomes

$$
\widehat{x}(k) + \sum_{i=1}^{n}\widehat{\theta}_i x(k-i) = \widehat{\phi}_1 \delta_1(k) \tag{31}
$$

The binary signal $\delta_1(k)$ is either 1 or -1 that is almost identical to the output of the equalizer shown in Eq. (3) with its coefficients computed by minimizing the cost function defined in equation (4). The main difference is that the coefficients $\theta_i,\ldots,\theta_n$ together with ϕ_1 in equation (31) are obtained by minimizing a global cost function rather than a local cost function at each time step k.

There are a total of N constant gain coefficients $\widehat{\phi}_1,\ldots,\widehat{\phi}_N$ for computing the output of the blind equalizer in Eq. (30), where $|\ \widehat{\phi}_i\ |$ is the i-th modulus. The open-loop prediction, Eq. (30), may thus be called the *adaptive multi-modulus blind predictor* or more precisely the *adaptive N-modulus blind predictor*. In [15], the multiple modulus concept has been proposed to deal with blind equalization of signal, such as M-ary QAM (quadrature amplitude modulation). The proposed approach in this paper differs from the approaches in [15, 16]

by not presetting the modulus. Furthermore, our proposed approach is a combination of recursive least-squares and the Decision Directed (DD) [9, 10] equalizer. In [15] the first proposed algorithm, the MMA (multiple modulus algorithm), uses a straightforward generalization of the cost function for the Constant Modulus Algorithm (CMA) [7, 8] to derive its update. The second one, DAMA (decision adjusted modulus algorithm), is a hybrid of the CMA and the DD approaches.

2.3.2 Closed-loop prediction for speech coding

In speech coding, the objective is to use the least number of bits in the digital representation of the speech signal, that is, to compress the signal $x(1), x(2), \ldots, x(\ell)$ where ℓ is the length of the signal. The open-loop prediction cannot be used for such purpose. It can only be done by the closed-loop prediction equation, i.e.,

$$\widehat{x}(k) = -\sum_{i=1}^{n} \widehat{\theta}_i \widehat{x}(k-i) - \sum_{i=1}^{N} \widehat{\phi}_i u_i(k); \quad k = 1, \ldots, \ell \tag{32}$$

where the initial signal $\widehat{x}(0), \widehat{x}(-1), \ldots, \widehat{x}(-n+1)$ are set to zero. For a better prediction, one may shift the starting prediction point from $k = 1$ to $k = n+1$ and set the first n points to be the actual signal. However, it will increase few bits for the additional n initial points to be stored and transmitted for speech coding.

From equation (32), it is clear that the speech signal can be represented by the n equalizer coefficients $\widehat{\theta}_1, \ldots, \widehat{\theta}_n$, the N gain coefficients $\widehat{\phi}_1, \ldots, \widehat{\phi}_N$, and the ℓ N-bit binary input input sequences. Assuming that each parameter may be accurately quantized by an M-bit binary number. As a result, a total of $M(n+N) + N\ell$ bits plus the codebook will be able to represent the speech signal of length ℓ. If the sampling rate is 8 KHz, a resulting bit rate is $[M(n+N) + N\ell]/8000$ kb/s.

2.4 Variable frame of input signal

To reduce the amount of data that must be transmitted, the conventional coders usually deal with the speech by frames (blocks) of samples. However, this may not be the best way to describe the non-stationary behavior of the sound sources. On the other hand, the coding performed sample-by-sample always results in a larger number of data to be transmitted. But these precise coding based on sample-by-sample can always produce high quality with negligible coding distortion and delay. The blind adaptive prediction was
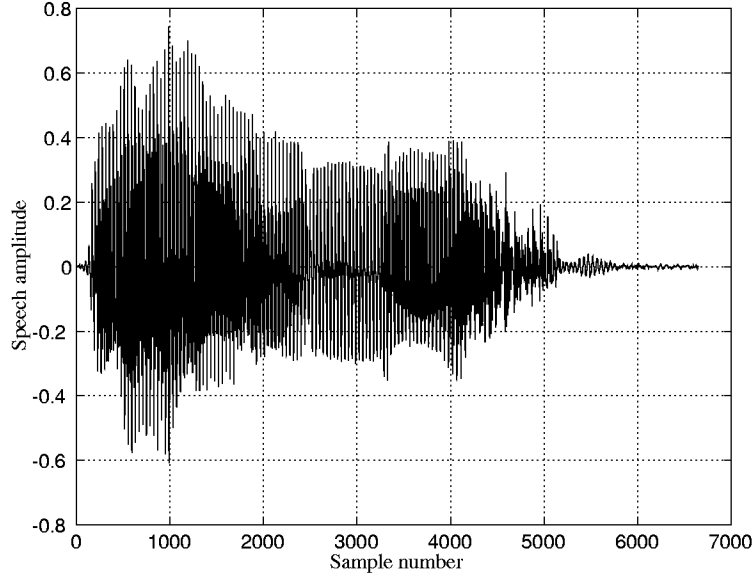
14

Figure 3: Original speech signal 'while new'

originally proposed to carry out the prediction on the sample-by-sample basis to overcome the non-stationary problems.

Here we introduce a variable frame concept for transmission and storage of input signal. The sample-by-sample scheme is used to do precise coding and obtain the high quality. Then, the input (codevector) sequence $\mathbf{\Delta}(k)$ for $k = 1, \ldots, \ell$ is blocked into frames to be transmitted or stored. Each frame contains an identical index from the binary codebook. The length of each frame is varying. Only the beginning point and the length of the frame need to be stored and transmitted. Hence, the resulting number of bits to represent the speech signal can be reduced. It may greatly reduce the bit rate for transmitting or the space for storing the input signal.

3 SIMULATION

The simulation is performed using a 10-th order equalizer for a speech signal shown in figure 3 where the total sample number is 6650. Several binary codebooks of different bit numbers are generated for searching the optimum input sequence. We set the forgetting factor $\lambda = 0.999$ and the initial value of $\mathbf{P}(0) = 1000\mathbf{I}_{11}$. First, the distribution of the

15
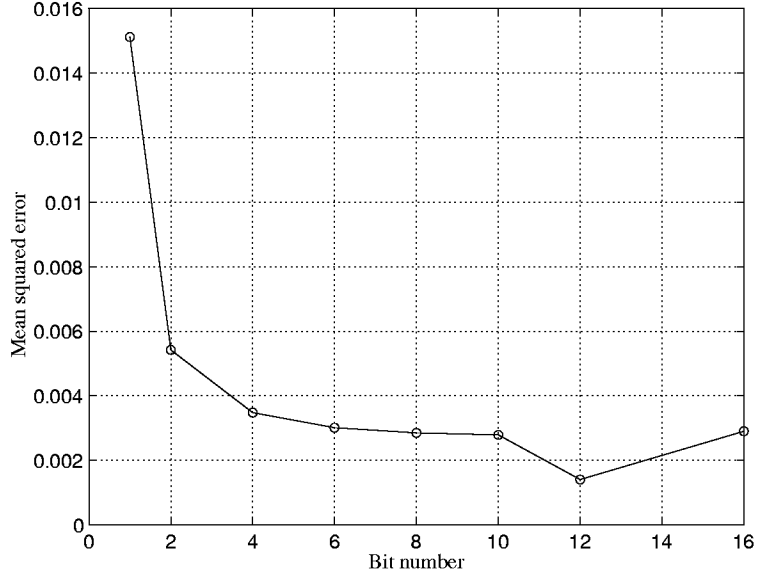
Figure 4: Mean squared error via different bit number codebooks

reconstructed error with various bit codebooks is examined. Figure 4 shows the plot of the mean squared error between the synthesized signal $\hat{x}(k)$ and its original signal $x(k)$, i.e., $\sum_{k=1}^{6650}(x(k) - \hat{x}(k))^2/6650$. When the bit number increases, there is a significant decrease in the mean squared error. However, it also leads to a considerable computational cost because of the search in an expanded codebook. As a 10-bit Gaussian codebook is adopted in CELP [3], we also focus on the bit number less than 10. This is due to the fact that for those codebooks with bit number more than 10 (where a 10-bit codebook consists of $2^{10} = 1024$ codevectors), the number of codevectors increases exponentially with the number of bits resulting in a high computational cost while searching for the optimum codevector. Let us now examine the results using a 4-bit codebook and a 8-bit codebook. For the 4-bit codebook, the gain vector for the input, $\Phi(k)$, is plotted in figure 5, and the coefficients of the equalizer is shown in figure 6. The synthesized speech generated by the closed-loop prediction can be found in figure 7. Figure 8 is the error between the original signal and the synthesized one, simply by subtraction. Since the recursive algorithm involves the coefficient smoothing, all parameters converge to a constant value. The gain vector approaches a constant value after about 2000 samples. These well-behavior gains indicate that the number of bits employed to do coding is sufficient. The synthesized speech signal

16
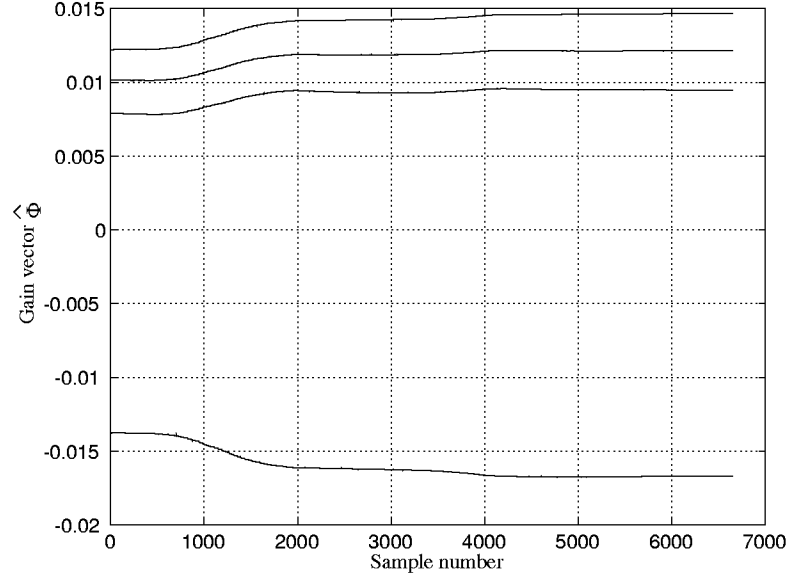
Figure 5: The gain vector for 4-bit coding

turns out to retain a good quality as expected. While listening to the sound, despite of the existing error, listeners can still hear it clearly.

The results obtained from the 8-bit codebook are shown in figures 9, 10, 11, and 12. The adapted gain vector is shown in figure 9, and the coefficients of the equalizer can be found in figure 10. The synthesized speech signal is in figure 11. In figure 12, we show the error between the original speech and synthesized one. The primary means of measuring performance is through subjective testing by listeners. Consistent with the results in figures 8 and 12, the quality of 8-bit coding is better than that of 4-bit coding. It is hard to distinguish audible differences the original and synthesized signal using 8-bit coding. It means that the quality of synthesized speech is reliable.

Revisiting the result in figure 4 for the bit number less than 10, the mean squared errors are, in fact, slightly different except those from 1-bit and 2-bit. Let us compare and discuss these cases in a detail manner. From the geometric theory as discussed in the previous section, the vector $\mathbf{X}$ must be representable in terms of the space generated by the columns of $\mathbf{V}^T$ to obtain an optimal solution for $\boldsymbol{\Psi}$ that minimizes the error between the original speech and its synthesized one. The increase of N-bit representation is meant to expand the dimension of the space generated by $\tilde{\mathbf{X}}$. Theoretically, the larger the N-bit

17
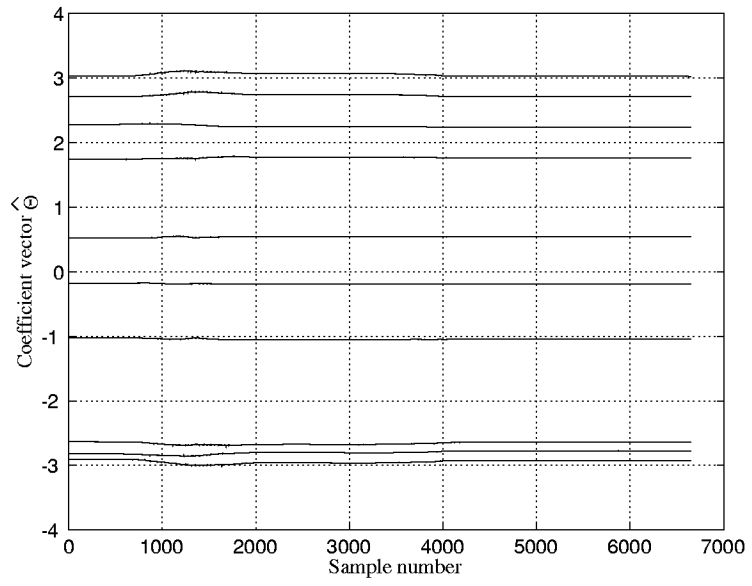
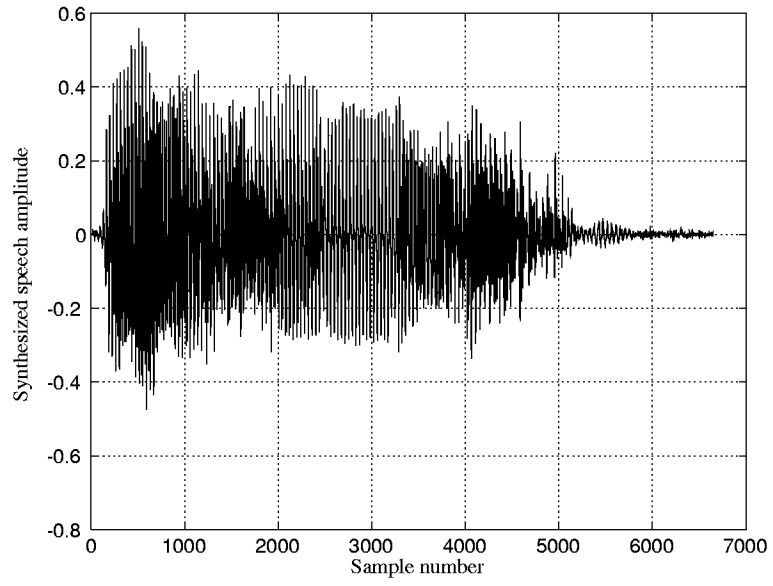Figure 6: The coefficients of the equalizer for 4-bit coding

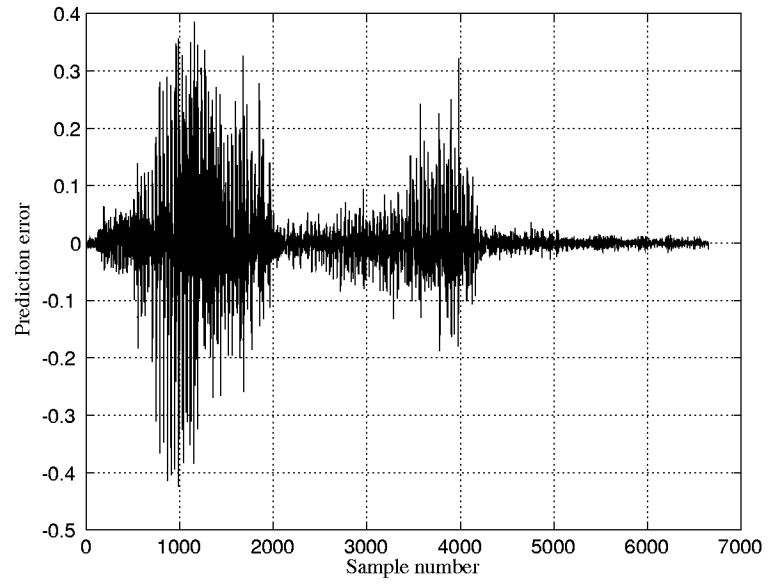Figure 7: The synthesized signal for 4-bit coding

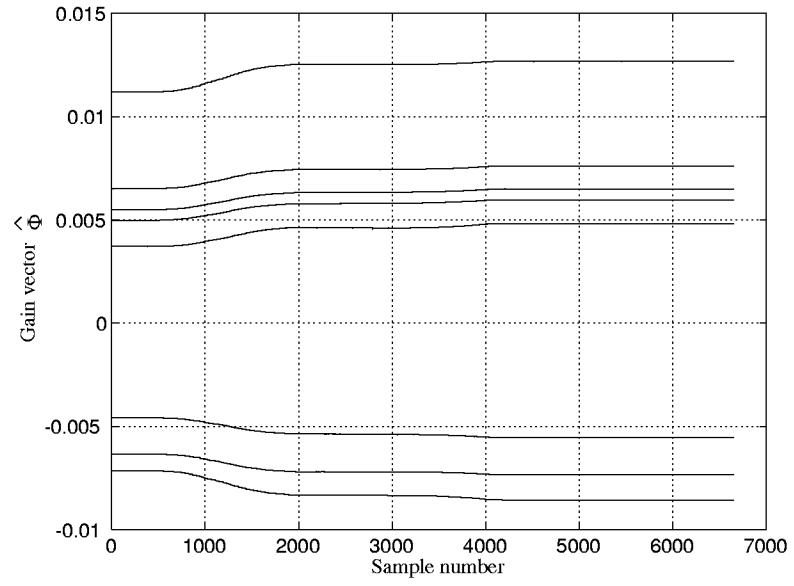Figure 8: The prediction error for 4-bit coding

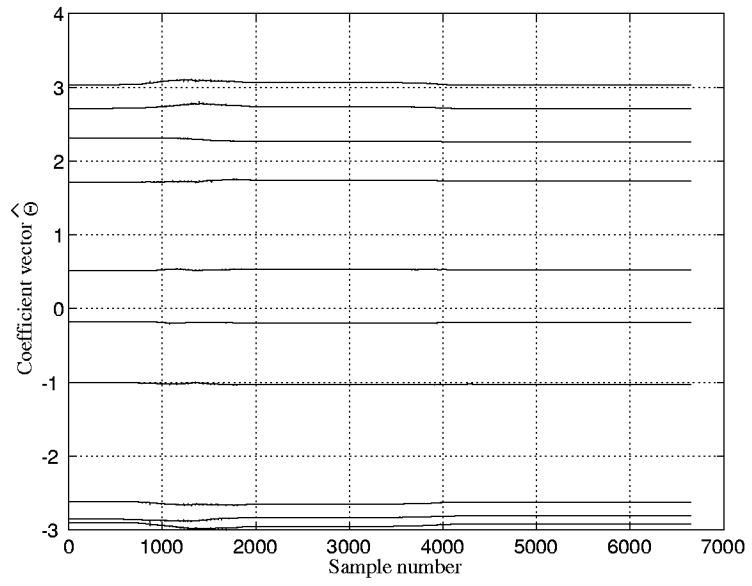Figure 9: The gain vector for 8-bit coding

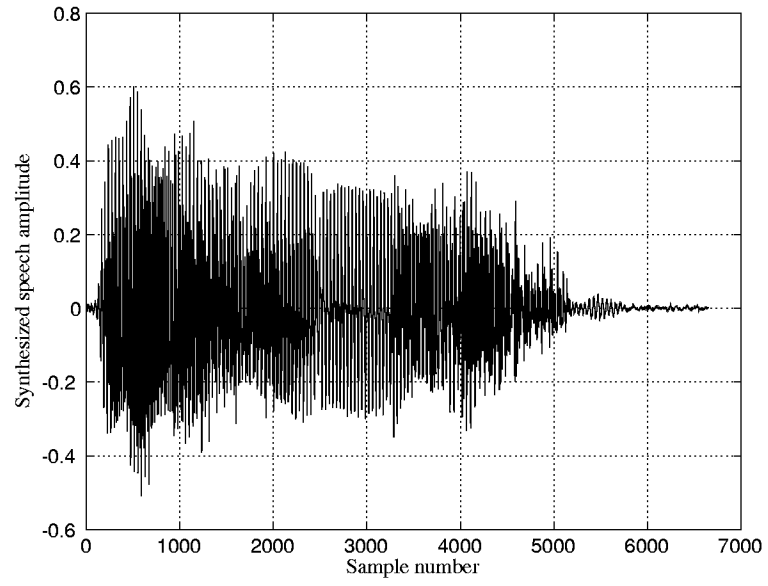Figure 10: The coefficients of the equalizer for 8-bit coding

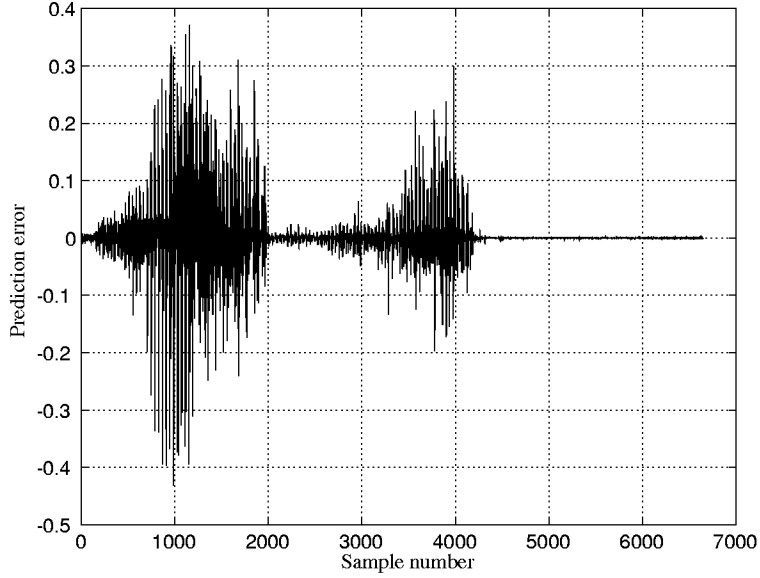Figure 11: The synthesized signal for 8-bit coding

Figure 12: The prediction error for 8-bit coding

representation is the better the optimal solution should be. However, the computational cost is also increased. An optimal N-bit representation depends on the desired quality of the synthesized signal, computational cost, and the data transmission rate. Let us compare and discuss the gain vector for the input signal for several different cases. Recall the gain vectors Φ shown in figures 5 and 9, and plot the gain vector results from a 1-bit codebook, a 2-bit codebook, and a 6-bit codebook in figures 13, 14, and 15, respectively. Clearly the 1-bit or 2-bit representation is not sufficient to produce a converged solution for the gain vector, that is, the space generated by V is not enough and the solution is poor. For a 4-bit representation, the solution improves considerably. Cases with 6-bit, 8-bit, or even 10-bit, show that the results are approximately the same but the extra bits enhance the quality of synthesis. Beyond 10-bit, the quality enhances slightly, but the computational cost increases considerably for just a bit added.

In the following, we introduce an improved scheme to advance the limited quality without significantly increasing the cost of computation. As used in the differential pulse code modulation (DPCM) [26], speech quality in LPC can be improved at the expense of a higher bit rate by computing and transmitting a residual error. We employ a low-bit codebook to encode both the signal and its residual. A 4-bit codebook as discussed earlier is sufficient
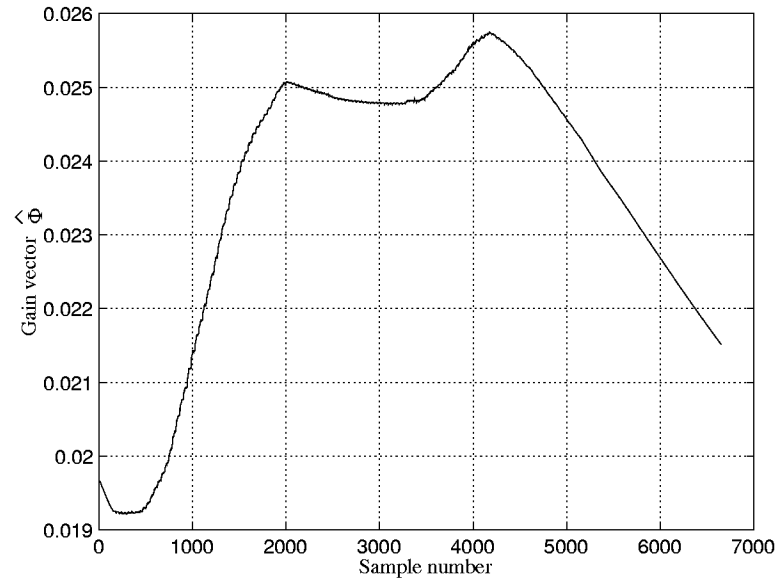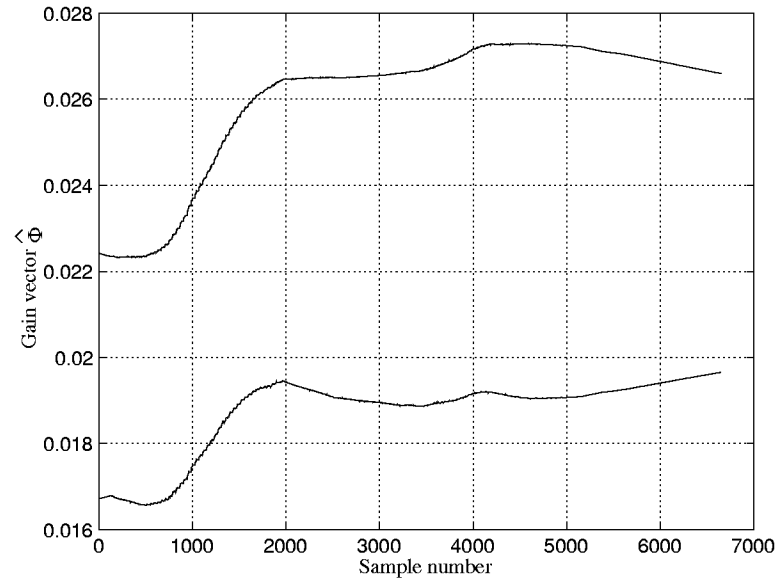
21

Figure 13: The gain vector for 1-bit coding

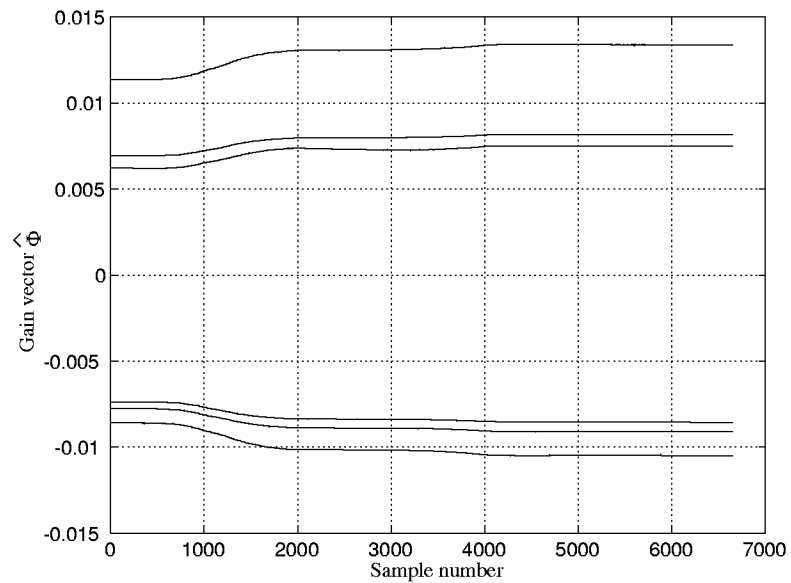Figure 14: The gain vector for 2-bit coding

22

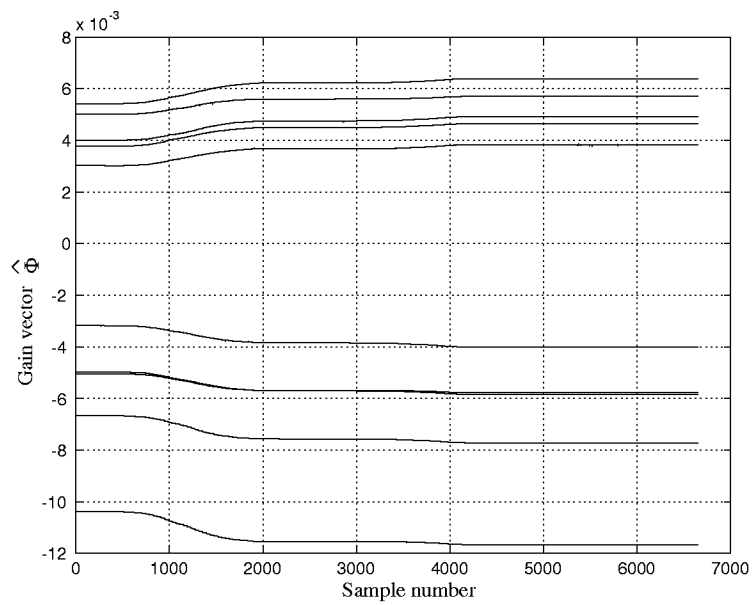Figure 15: The gain vector for 6-bit coding

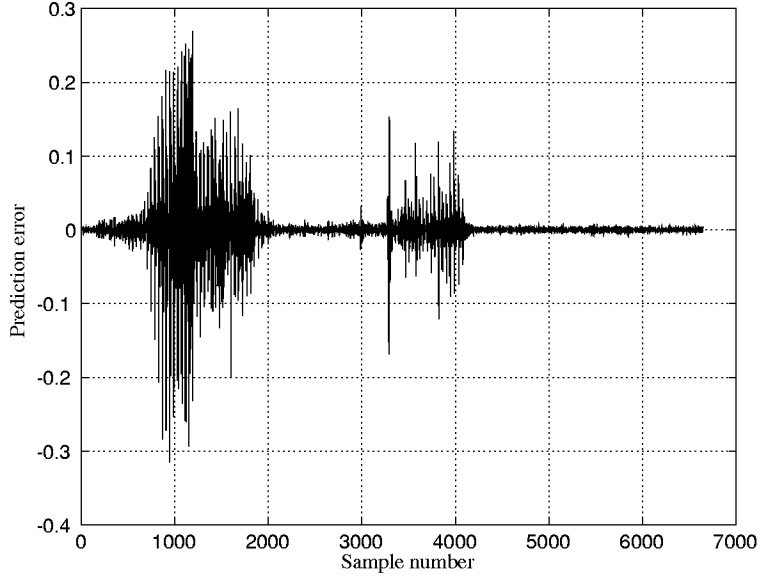

Figure 16: The gain vector for 10-bit coding

23

Figure 17: The prediction error for 8-bit coding by improved scheme

to obtain a good solution for the gain vector of input signal. As shown in figure 8, the error is encoded again by the same 4-bit binary codebook. Figure 17 shows that the error of the synthesis is reduced when compared to that in figure 12. Moreover, when listening to them, the quality of the improved scheme is better than just encoding the signal by a 8-bit codebook.

In this improved scheme, we repeat the encoding process of the prediction residual and thus the computational cost of encoding is doubled. Since the proposed algorithm is computationally efficient, this cost increase is still acceptable. Because we split a 2N-bit coding to two N-bit coding, the codebook for optimal search is reduced by 2^N in dimension. The computational cost for tow N-bit coding is still much cheaper than the 2N-bit coding. Moreover, it was found that beyond certain number of bits in the codebook, the quality of synthesized signal does not improve. As the results shown in figure 5, the gain of the input is as stable as that in figure 9, implying that the space spanned by $\mathbf{X}$ with 4-bit coding is sufficient enough to produce a good solution of $\widehat{\boldsymbol{\Psi}}$. By applying the multi-step coding, the quality by two 4-bit coding is considerably improved than that by one 8-bit coding. This improved scheme not only enhances the quality of speech as the same 8 bits stored or transmitted, but also decrease the computational cost during the search for an optimum
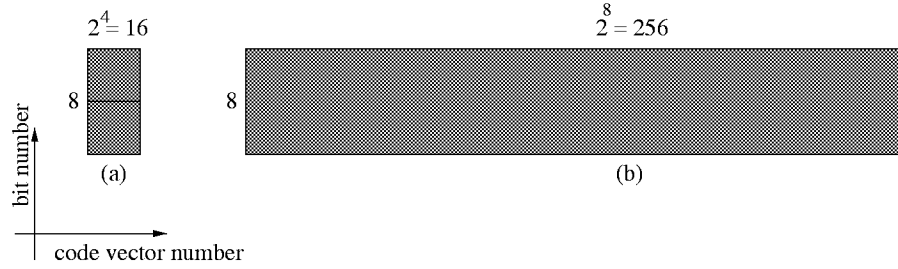
Figure 18: Codebooks for searching an optimum input sequence (a) improved scheme (b) original shceme

8-bit input sequence. Figure 18 shows the sizes of 4-bit and 8-bit codebooks.

4 CONCLUSIONS

In this report, a new method is developed and implemented for speech coding and synthesis. The conventional linear predictive coding requires two computational steps, i.e., coefficient estimation of an all-pole model, and quantization of the prediction residual. The model coefficients are estimated by minimizing the mean squared error. The prediction residual is quantized to be used as the input signal during the process of speech coding and signal synthesis. On the other hand, the blind adaptive predictive coding estimates the coefficients of the blind equalizer with the assumption that the output of the equalizer (i.e., the input for speech coding) is a priori fixed at uncertain values. It generally produces a poor estimate of the coefficients and a poor quality of the signal synthesis.

In contrast, the proposed method uses the deterministic approach to simultaneously estimate the combined coefficients of the blind equalizer and the binary input excitation. The linear geometric theory is used to establish the theoretical background of the proposed method. The combined coefficients are estimated by minimizing the angle between a vector in the direction of the signal and the space generated by the shifted signal and the binary input. A recursive algorithm is introduced for computational ease and real-time implementation. Simulations have shown that the quality of the synthesized signal can be significantly improved from 1-bit coding to 4-bit coding. Beyond 4-bit coding, the quality enhances slightly but the computational cost increases considerably. To overcome this problem, the encoding process is repeated on the prediction residual using the same 4-bit

codebook. With much less computational cost, the repetitive process produces the quality of the synthesized signal with the 4-bit codebook better than that with the 8-bit coding.

The proposed technique provides a totally different framework for voice coders. The concept based on the linear geometric theory gives a new direction to explore more fundamental works and applications.

References

[1] J. R. Deller, JR., J. G. Proakis, and J. H. L. Hansen, *Discrete-time Processing of Speech Signal*, Prentice-Hall, 1993.

[2] W. B. Kleijn and K. K. Paliwal, *Speech Coding and Synthesis*, Elsevier science B. V., 1995.

[3] M. R. Schroeder and B. S. Atal, "Code-excited Linear Prediction (CELP): High-quality Speech at Very Low Bit Rates," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 937-940, 1985.

[4] I. A. Gerson and M. A. Jasiuk, "Vector Sum Excited Linear Prediction (VSELP) Speech Coding at 8kbps," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Vol. 1, pp. 461-464, 1990.

[5] H. G. Fehn and P. Noll, "Multipath Search Coding of Stationary Signals with Applications to Speech," *IEEE Trans. Communications*, Vol. COM-30, pp. 687-701, 1982.

[6] A. Buzo, A. H. Gray, Jr., R. M. Gray, and J. D. Markel, "Speech Coding Based upon Vector Quantization," *IEEE Trans. Acoust., Speech, Signal Proc.*, Vol. ASSP-28, pp. 562-574, 1980.

[7] D. Godard, "Self-recovering Equalization and Carrier Tracking in Two-dimensional Data Communication Systems," *IEEE Trans. Commun.*, Vol. COM-28, pp. 1867–1875, Nov. 1980.

[8] A. Benveniste and M. Goursat, "Blind Equalizers," *IEEE Trans. Commun.*, Vol. COM-32, pp. 871–882, Aug. 1982.

[9] J. Proakis, "Advances in Equalization for Intersymbol Interference," *Advances in Communications Systems*, Vol. 4, Academic Press, 1975.

[10] O. Macchi and E. Eweda, "Convergence Analysis of Self Adaptive Equalizers," *IEEE Trans. on Information Theory*, Vol. IT-30, No. 2, Mar. 1984.

[11] S. Haykin, *Blind Deconvolution*, Englewood Cliffs, NJ:Prentice-Hall, 1994.

[12] Y. C. Chen, T. Stathaki and A. G. Constantinides, "Adaptive Prediction Based on Blind Equalisation Principles," *ISMIP-96*, Taiwan, pp. 301–307, Dec. 1996.

[13] ——————-, "Nonlinear Adaptive Prediction of Nonstationary Signals with Application to Speech Coding," *IEEE Signal Processing, MMSP-97*, Princeton, New Jersey, June 1997.

[14] Y. C. Chen , J.-N. Juang, "A Novel Approach for Adaptive Signal Processing," *NASA/CR-1998-206920 ICASE Report No. 98-11*, Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, VA 23681-2199, February 1998.

[15] W. Sethares, G. Rey, and C. Johnson, Jr., "Approaches to Blind Equalization of Signals with Multiple Modulus," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 972-975, 1989.

[16] R. Axford, Jr., L. Milstein, and J. Zeidler, "The Transversal and Lattice Multiple Modulus Algorithms for Blind Equalization of QAM Signals," *MILCOM IEEE Military Communication Conference*, pp. 586-591, 1995.

[17] Jer-Nan Juang, *Applied System Identification*, Englewood Cliffs, New Jersey 07632, 1994.

[18] B. S. Atal, "Predictive Coding of Speech at Low Bit Rates," *IEEE Trans. Communications*, Vol. COM-30, pp. 600-614, 1982.

[19] F. Itakura and S. Saito, "Analysis Synthesis Telephony Based on the Maximum Likelihood Method," *Proc. 6th Int. Cong. Acoust.*, C-5-5, 1968.

[20] B. S. Atal and M. R. Schroeder, "Predictive Coding of Speech Signals," *Proc. 6th Int. Cong. Acoust.*, C-5-4, 1968.

[21] F. Sadaoki, *Speech Processing Systems*, 2nd ed. Marcel Dekker, 1989.

[22] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Prentice-Hall, Englewood Cliffs, NJ, 1978.

[23] J. R. Deller, Jr., J. G. Proakis and J. H. L. Hansen, *Discrete-Time Processing of Speech Signals*, Prentice-Hall, Inc., NJ., 1987.

[24] S. Kay, *Modern Spectral Estimation: Theory and Application*, Englewood Cliffs, NJ:Prentice-Hall, 1988.

[25] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ:Prentice-Hall, 1996.

[26] N. S. Jayant, "Digital Coding of Speech Waveforms:PCM, DPCM, and DM Quantizers," *Proc. of IEEE*, Vol. 62, pp. 611-632, May 1974.

[27] G. J. Bierman, *Factorisation Methods for Discrete Sequential Estimation*, Academic Press, 1977.

[28] K. J. Astrom, and B. Wittenmark, *Adaptive Control*, Second Edition, Addison-Wesley Publishing Company, 1995.

			Form Approved
REPORT DOCUMENTATION PAGE			*Form Approved* *OMB No. 0704–0188*

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE October 1999	3. REPORT TYPE AND DATES COVERED Technical Memorandum

4. TITLE AND SUBTITLE
Signal Prediction With Input Identification

5. FUNDING NUMBERS
WU 632-10-14-04

6. AUTHOR(S)
Jer-Nan Juang and Ya-Chin Chen

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NASA Langley Research Center
Hampton, VA 23681–2199

8. PERFORMING ORGANIZATION REPORT NUMBER
L–17889

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Washington, DC 20546–0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
NASA/TM–1999–209705

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified-Unlimited
Subject Category 39   Distribution: Standard
Availability: NASA CASI (301) 621–0390

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
A novel coding technique is presented for signal prediction with applications including speech coding, system identification, and estimation of input excitation. The approach is based on the blind equalization method for speech signal processing in conjunction with the geometric subspace projection theory to formulate the basic prediction equation. The speech-coding problem is often divided into two parts, a linear prediction model and excitation input. The parameter coefficients of the linear predictor and the input excitation are solved simultaneously and recursively by a conventional recursive least-squares algorithm. The excitation input is computed by coding all possible outcomes into a binary codebook. The coefficients of the linear predictor and excitation, and the index of the codebook can then be used to represent the signal. In addition, a variable-frame concept is proposed to block the same excitation signal in sequence in order to reduce the storage size and increase the transmission rate. The results of this work can be easily extended to the problem of disturbance identification. The basic principles are outlined in this report and differences from other existing methods are discussed. Simulations are included to demonstrate the proposed method.

14. SUBJECT TERMS
Signal Processing, Speech Coding, Data Compression, Signal Prediction

15. NUMBER OF PAGES
33

16. PRICE CODE
A03

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT